TECHNICAL REPORT 65-5

H. KUSHNER

# STOCHASTIC STABILITY AND THE DESIGN OF FEED BACK CONTROLS

MAY 1965



CENTER FOR DYNAMICAL SYSTEMS

DIVISION OF APPLIED MATHEMATICS

BROWN UNIVERSITY

(1)

# STOCHASTIC STABILITY AND THE

# DESIGN OF FEEDBACK CONTROLS

H. Kushner †

† This research was supported in part by National Aeronautics and Space Administration under Grant No. NGR-40-002-015 and in part by the United States Air Force through the Air Force Office of Scientific Research under Grant No. AF-AFOSR-693-64.

Stochastic Stability and the

Design of Feedback Controls

H. Kushner

## 1. INTRODUCTION

The object of this paper is to describe the stochastic extensions of the various techniques for using the second method of Liapunov to aid the construction and analysis of feedback controls [1-8]. The method appears to be useful for design and analysis, although it is too early to make a final judgment. Much depends on future success in finding suitable Liapunov functions, and understanding the relationship between the loss function and the desired behavior of the control system.

The deterministic methods have been motivated by considerations of the following nature. Consider the optimal control problem with control $u$, and system

$$\dot{x} = f(x, u)$$

and cost

$$C^u(x) = \int_0^\tau k(x, u)dt, \quad C^u(\partial S) = 0$$

where $\tau$ is the time of contact with $\partial S$, the boundary of a target set $S$. The minimum cost

$$\underline{C}(x) = \min_u C^u(x),$$

is achieved by $u = w$. If $\underline{C}(x)$ is sufficiently differentiable, then the Hamilton-Jacobi equation

$$d\underline{C}(x)/dt = \underline{C}'_x(x)f(x, w) = -k(x, w)$$

is satisfied (where $'$ is transpose and $\underline{C}_x$ is the gradient of $C$), and $w$ is the $u$ minimizing

$$[\underline{C}'_x(x)f(x, u) + k(x, u)].$$

In lieu of attempting to solve this problem, an alternative procedure has suggested itself to many authors (e.g.[1]-[8]). Choose a Liapunov function $V(x)$, and some $u(x)$ so that the system has suitable stability properties, and compute ($X$ is the state space)

$$V'_x(x)f(x, u) = - k_1(x, u).$$

where $k_1(x, u) \geq 0$ in $X + \partial S$. $X$ is the state space.

A comparison of $k_1(x, u)$ and $k(x, u)$ can yield useful information; e.g., whether $V(x)$ is greater or less than $\underline{C}(x)$, or stability properties of the controlled system, the nature of the problem for which $V(x)$ and $u$ are optimum, and whether some other calculatable control would minimize the cost $C^u(x)$, etc.

Similar results are achievable in the stochastic situation. Stochastic stability seems to be a more complicated subject than its deterministic counterpart, since the corresponding Liapunov functions do not decrease monotonically for each sample function. The effect of controls on the statistical behavior of the system can be made rather explicit in terms of a reduction of a bound on the probability of arbitrary deviations in the sample paths before hitting $\partial S$.

In part II several comparison and optimality theorems are proved. In part III the theorems are applied to the problem of choosing and analyzing the effect of feedback controls for several stochastic systems.

## 2. THE SYSTEM TO BE CONTROLLED

The object to be controlled is represented by the vector stochastic differential (Ito) equation

$$dx = f(x,\ u)dt + \sigma(x,\ u)dz, \tag{1a}$$

by which is meant (using the Ito [9] interpretation of the stochastic integral)

$$x_t = x_o + \int_o^t f(x_t,\ u(x_t))dt + \int_o^t \sigma(x_t,\ u(x_t))dz_t. \tag{1b}$$

$z$ is a vector Wiener process with independent components, $\dot{z}$ is commonly called white Gaussian noise;

$$\dot{x} = f(x,\ u) + \sigma(x,\ u)\dot{z}\ . \tag{1c}$$

$f$ is a vector with components $f_i$, and $\sigma$ is a matrix with components $\sigma_{ij}$. The process $x_t$ is confined to $X$.

Without the control parameter $u$, the meaning of (1b) and the conditions under which a solution (a stochastic process) exists and is unique is discussed in [9], [10]. To be secure in the mathematical development we assume these conditions. Let $\| \ \|$ be the Euclidean norm. For some finite positive K, let

$$\|f(x + \alpha, u + \beta) - f(x, u)\| \leq K\|\alpha\| + K\|\beta\|$$

(2)
$$\|\sigma(x + \alpha, u + \beta) - \sigma(x, u)\| \leq K\|\alpha\| + K\|\beta\|$$

$$\|f(x, u)\| \leq K[1 + \|x\|^2 + \|u\|^2]^{\frac{1}{2}}$$

(3)
$$\|u(x + \alpha) - u(x)\| \leq K\|\alpha\|$$

A control satisfying (3) is termed <u>admissible</u>. (3) implies continuity of $u(x)$. Note that $u = \text{sign } x$ is not admissible. Since the $K$ in (3) can be large, admissibility is probably not a serious restriction.

In certain cases, our results are valid if (2) and (3) replaced by local Lipshitz conditions. This is the case when the trajectories have appropriate stability properties (e.g., when the origin is stable w.p.1 in the sense of [15]).

The primary attractions of the model (1) are that it represents a rather large class of Markov processes with continuous sample paths, there is a large body of theory concerning it, and it seems that many physical problems can be modelled by it. The question of modelling will not be discussed. The identification of particular forms of (1) with particular physical problems is still an open problem in general (especially in the non-linear case). (Some interesting results in [11] clarify some of the questions of modelling.)

For each integer $r$, define the stochastic process

(4)
$$x^r_{n+1} = x^r_n + f(x^r_n, u(x^r_n))\Delta + \sigma(x^r_n, u(x^r_n))\delta z_n,$$

where $\delta z_n = z((n+1)\Delta) - z(n\Delta)$, and define $x^r(t) = x^r_n$ in the interval $(n+1)A > t \geq n\Delta$. Then, for a suitable sequence of $\Delta \to 0$, we have $x^r(t) \to x(t)$ with probability one for each $t$, where $x(t)$ is the solution to (1).

Some facts, to be used later, will be quoted. If $u(x)$ is admissible, $x_t$ is continuous with probability one, and is a Markov process; i.e., for any measurable set $A$ in $X$,

$$(5) \qquad P[x_{t+s} \in A \mid x_\sigma, \ \sigma \leqq t] = P[x_{t+s} \in A \mid x_t],$$

where the bar $|$ denotes conditional probability. A major, relatively recent, development in probability theory is the analysis and extensive use of the concept of random time (see [10], [12], [13] for details). An example of a random time is the first time that $x_t$ leaves an open set $A$; $\tau = \min \{t: x_t \notin A\}$. $\tau$ is a random variable. Loosely speaking, whether or not the event $\{\tau < t\}$ has occured (in the example, whether $x_s$ has left $A$ by time $t$) can be determined by observations on the $x_s$ process up to and including time $t$. (The set $\{\tau < t\}$ is in the $\sigma$-field determined by $x_s$, $s \leqq t$.)

The significance to control applications, of the concept of random time, will be seen in the sequel. If the process $x_t$ is confined to a set $X$ which is compact, and if $u$ is admissible, the process $x_t$ is, in fact, a strong Markov process. A strong Markov process has the Markovian property relative to random times. Let $\tau$ be a random time, then

$$P[x_{\tau+s} \in A \mid x_\delta, \ \delta \leqq \tau] = P[x_{\tau+s} \in A \mid x_\tau].$$

For example, let $x_t$ start in an open set $B$, let $\tau$ be the least time of leaving $B$, then for any non random $s$ the probability that $x_{\tau+s} \in A$ given $x_\tau$ and the paths up to $\tau$ equals the probability given only $x_\tau$. The strong Markov property is proved in [10].

## 3. THE CONTROL PROBLEM

The process $x_t$ is defined in a set $X$ in a Euclidean space. There is a set $S$ in $X$ given, and the main object of the control is to transfer $x_o = x$ to $\partial S$ in finite average time. In certain cases, infinite average times will be allowed. The proofs of the theorems we require assume that $X$ is compact (e.g., the proof of (8) for the operator $L^u$). This does not seem to be a restriction from the practical point of view since $X$ may be as large as desired. We may stop the process upon leaving some very large set, and estimate the probability of this event by (10). Also, to each $u$ and initial point $x_o = x$, there is the associated cost

$$C^u(x) = E_x^u \int_o^{\tau_u} k(x_t, u_t)dt \tag{6}$$

$E_x^u$ is the expectation and $\tau_u$ is the random time of arrival at $\partial S$, (provided that it is defined) and $k(x, u)$ is continuous and non-negative, and is referred to as the loss.

Define $\underline{C}(x) = \min_u C^u(x)$, provided that the minimizing $u$ is admissible. Part of the control problem is the comparison of $C^u(x)$, and $E_x^u \tau_u$ for various controls. Various restrictions may be placed on the control; it may be bounded, or its functional form may be restricted; e.g., it may be allowed to be a function of some, but not all, components of $x$. Some stability properties may also be of interest; e.g., an estimate of the probability that $x_t$ ever leaves some set $X'$, if $x_o$ is in $X'$, or some other qualitative information on the random paths.

A number of relevant results and examples on stability are in [14]-[16]. [17] is concerned with ergodic properties of the processes and utilizes certain properties of stochastic Liapunov functions.

## 4. OTHER MATHEMATICAL PRELIMINARIES

Let $u$ be admissible. Define the operator

$$L^u = \frac{\partial}{\partial t} + \sum_i f_i(x, u) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j} S_{ij}(x, u) \frac{\partial^2}{\partial x_i \partial x_j} \tag{7}$$

$$S_{ij} = \sum_k \sigma_{ik} \sigma_{jk}$$

$L^u$ is the differential generator of the $x_t$ process, with control $u$. We say that $V(x)$ is in the domain of $L^u$ ($V(x) \in D(L)$) if $V(x)$ is a non-negative, scalar valued function with continuous second derivatives and the sets $\{x : V(x) \leq c\}$ are compact and connected,† for all $c$ less than some $c_o > 0$. Such a $V(x)$ will also be called a Liapunov function, or a Liapunov function in a region $R$, if $L^u V \leq 0$ in $R$ for the given $u$. Note that $L^u V(x) = dV(x)/dt = V_x'(x) f(x, u)$ in the deterministic case.

Since $X$ is compact and $x_t$ is a strong Markov process, Dynkin's formula [10]

$$E_x^u V(x_\tau) - V(x) = E_x^u \int_0^\tau L^u V(x_s) ds \tag{8}$$

holds for all random times $\tau$ with $E_x^u \tau < \infty$. (8) underlies many of the results of the sequel. It says, in effect, that $V(x)$ is the average value of the integral of the 'stochastic derivative' $L^u V(x)$. The compactness of $X$ and the finiteness of $E_x^u \tau$ are important in establishing its validity. In other cases the operator $L^u$, for which (8) is valid, will be an extension of (7), but this is beyond our purpose.

Let $V(x)$ be in $D(L^u)$ in the region $R \in X$,

---

†The domain of $L^n$ is obviously larger than our $D(L)$, but $D(L)$ suffers less.

$$R = \{x : V(x) < \lambda\} - \{x : V(x) \leq \lambda_o\}, \quad \lambda > \lambda_o$$
$$\partial R = \{x : V(x) = \lambda, \quad V(x) = \lambda_o\}.$$

Let $L^u V < 0$ in $R$ and $L^u V(x) \leq 0$ on $\partial R$. Let $\tau_u$ be the random time to $\partial R$, starting at $x \in R$. It can be proved, using the continuity of $x_t$ and $L^u V(x)$, and the compactness of $X$, that $\tau_u$ will exist (although $E_x^{\tau_u}$ may not be finite), and that the integral in (8) is defined and finite, and that (8) is valid. (See [15]). Since $x_t$ is continuous with probability one, $x_t$ cannot leave $R$ without touching $\partial R$ (with probability one). We have

$$E_x^u V(x_\tau) = \lambda_o P[\sup_{\tau_u \geq t \geq 0} V(x_t) \leq \lambda_o] + \lambda P[\sup_{\tau_u \geq t \geq 0} V(x_t) \geq \lambda] . \tag{9}$$

Letting $\lambda_o = 0$, and noting that the integral in (8) is non-positive in $R + \partial R$,

$$P[\sup_{\tau_u \geq t \geq 0} V(x_t) \geq \lambda] = (V(x) - E_x^u \int_o^{\tau_u} L^u V(x) dt)/\lambda \leq V(x)/\lambda. \tag{10}$$

(10) can also be derived by showing that $V(y_t)$ is a non-negative super martingale, where $y_t$ is the $x_t$ process stopped at time $\tau_u$; then the inequality (10) is the non-negative super martingale inequality. Both methods may be used when time is discrete. If $L^u V^n(x) \leq 0$ in $R$ for any real number $n \geq 1$, then

$$P\left[\sup_{\tau_u \geq t \geq 0} V(x_t) \geq \lambda\right] \leq V^n(x)/\lambda^n \tag{11}$$

which is an improvement over (10).

In general, we will try to improve (10) by finding the maximum $n$ for which $L^u V^n(x) \leq 0$ in $R$. This method is not generally the best for obtaining probability bounds on the behavior of components of $x$.

## Part II. COMPARISON AND OPTIMALITY THEOREMS

It is always assumed that $X$ is compact, $k(x, u) \geq 0$ and continuous, and that (2), (3) are satisfied. The purpose of the theorems is to allow a comparison of the costs and stability properties resulting from the use of different controls, and to obtain upper and lower bounds on $\underline{C}(x)$ without actually solving the minimization problem. The symbols $\tau_o$, $\tau_u$ are the random times to transfer $x_o = x$ (in some given initial set) to $\partial S$, the boundary of the target set $S$, in the cases of no control, and control $u$, respectively.

$\partial R$, The theorems use the assumption $E_x^u \tau_u < \infty$. When $L^u V(x) < 0$ in $X - S$ and $L^u V(x) \leq 0$ on $\partial S$ and does not depend on time, the finiteness assumption may be dropped. The modification will be used occasionally in the examples. It is usually of little consequence, since a slight enlargement of the target set will usually assure that $E_x^u \tau_u < \infty$.

### Theorem 1

Assume that there is an optimal admissible control $w$ with $E_x^w \tau_w < \infty$. Let $u$ be admissible and $E_x^u \tau_u < \infty$. Let $V_1(x)$ be in $D(L)$ and $V_1(\partial S) = 0$, and

$$L^u V_1(x) + k(x, u) < 0 \tag{14}$$

*in* X - S. *Then*

(15) $$V_1(x) > \underline{C}(x).$$

Also, for any $\lambda > 0$,

(16) $$P[\sup_{\tau_u \geq t \geq 0} V_1(x_t) \geq \lambda] \leq EV_1(x)/\lambda.$$

*If there is a* $V_2(x)$ *in* $D(L)$ *with* $V_2(\partial S) = 0$ *and, for all admissible* $u$,

(17) $$L^u V_2(x) + k(x, u) > 0,$$

*in* X - S, *then*

(18) $$V_2(x) < \underline{C}(x).$$

(In the event that there is a non-admissible control for which the problem has a meaning, and which minimizes $C^u(x)$, then the first part of the theorem still holds.)

*Proof*:

(8) may be applied to $V_1(x)$ and $\tau_u$. Thus,

$$V_1(x) - E_x^u V_1(x_{\tau_u}) > E_x^u \int_0^{\tau_u} k(x, u)dt \geq \underline{C}(x).$$

Since $E_x^u \tau_u < \infty$, $x_{\tau_u}$ is on $\partial S$ w.p.1. Since X is bounded and $V_1(\partial S) = 0$, we have $E_x^u V_1(x_{\tau_u}) = 0$, and (15) follows.

Since $w$ is admissible, and $V_2(x)$ is in $D(L)$ and $E_x^w \tau_w < \infty$, the application of (8) to $V_2(x)$ and $\tau_w$ yields

$$V_2(x) - E_x^w V_2(x_{\tau_w}) < E_x^w \int_0^{\tau_w} k(x, w)ds = \underline{C}(x),$$

$E_x^w V_2(x_{\tau_w}) = 0$ by a repetition of a former argument. (16) follows from (10).

Corollary 1

Let the optimal admissible control exist, and let $V(x)$ satisfy the conditions on $V_1(x)$. Let $u$ and $w$ be admissible controls, $E_x^w \tau_w < \infty$, $E_x^u \tau_u < \infty$, and, for all such $u$,

$$L^u V(x) + k(x, u) \geq 0$$

with equality when $u = w$. Then

$$V(x) = \underline{C}(x)$$

and $w$ is optimal.

Proof:

The statement follows from Theorem 1, by setting $V(x) = V_1(x) = V_2(x)$ and replacing all $>$ by $\geq$.

Remark:

If there is an admissible control which is optimal and a $V(x)$ satisfying the conditions of the corollary is available, then the corollary partially justifies the usual result of dynamic programming; i.e., that the optimum control minimizes (19) and that the solution of (19) is $V(x) = \underline{C}(x)$.

$$\min_u [L^u V(x) + k(x, u)] = 0 \tag{19}$$

Corollary 2

(10) is valid and $\tau_u$ is defined with probability one, when $k(x, u) > 0$ in $X - S$. Under this condition, the condition on definiteness of the average arrival times can be dropped, and we have a true optimality theorem, (the stochastic counterpart of the Hamilton-Jacobi equation theorem in [20]).

Theorem 2

Let

$$C^u(x) = E_x^u \int_0^{\tau_u} [k(x) + \ell(x, u)]dt$$

*where* $k \geq 0$, $\ell \geq 0$ and $\ell(x, 0) = 0$. *Let* $L^o$ *correspond to* $u = 0$. *Let* $E_x^o \tau_o < \infty$, $V(x)$ *in* $D(L)$ *and* $V(\partial S) = 0$ *and*

$$L^o V(x) + k(x) = 0.$$

*For some* $u$, *let* $E_x^u \tau_u < \infty$, *and*

$$L^u V(x) + k(x) + \ell(x, u) < 0. \tag{20}$$

*Then*

$$C^o(x) = E_x^o \int_o^{\tau_o} k(x)dt > E_x^u \int_o^{\tau_u} [k(x) + \ell(x, u)]dt = C^u(x). \tag{21}$$

*If* $\geq$ *replaces* $>$ *in* (20), *it does so in* (21).

*Proof:*

The proof is essentially that of Theorem 1. From (8)

$$E_x^u \int_o^{\tau_u} [L^u V(x) + k(x) + \ell(x, u)]dt < 0 = E_x^o \int_o^{\tau_o} [L^o V(x) + k(x)]dt \;,$$

$$- V(x) + E_x^u V(x_{\tau_u}) + C^u(x) < - V(x) + E_x^o V(x_{\tau_o}) + C^o(x).$$

Since $x_{\tau_u}$ and $x_{\tau_o}$ are on $\partial S$ w.p.1, and $V(x)$ is bounded in $X - S$, the theorem follows.

*Remark:*

Consider the special case

$$dx = f(x, u)dt + \sigma(x)dz,$$

where $\sigma$ does not depend on $u$, and where (with $V(\partial S) = 0$)

$$(22) \qquad k(x) = -L^oV(x) = -V_x'(x)f(x, 0) - \frac{1}{2}\sum_{i,j} \frac{\partial^2 V(x)}{\partial x_i \partial x_j} S_{ij}(x).$$

With $u \neq 0$,

$$(23) \qquad -L^uV(x) = -V_x'(x)f(x, u) - \frac{1}{2}\sum_{i,j} \frac{\partial^2 V(x)}{\partial x_i \partial x_j} S_{ij}(x).$$

By Theorem 2, for any $u$ such that

$$L^uV(x) - L^oV(x) + \ell(x, u) \leq 0$$

or

$$(24) \qquad V_x'(x)\,[f(x, u) - f(x, 0)] + \ell(x, u) < 0,$$

we have

$$C^u(x) < C^o(x).$$

Although the theorem states that a control will decrease the cost under certain conditions, accurate estimates of the decrease are usually difficult to obtain. Estimates of the effect of the control on the probability (16) are readily available (see the examples). We obtain the best improvement of the value of (16) with the $u$ which minimizes (20). Otherwise, the problem of selecting one, from among the many controls which may satisfy (24), is open.

Theorem 3 gives a condition under which $E_x^u \tau_u < \infty$ is assured.

Theorem 3

*Let* $V(x)$ *be in* $D(L)$. *If, for some* $\epsilon > 0$,

$$L^u V(x) = -k_1(x, u) \leq -\epsilon$$

*in* $X - S$, *then* $\tau_u$ *exists* w.p.1 *and*

$$E_x^u \tau_u \leq V(x)/\epsilon < \infty.$$

*Let there exist an optimum* $w$ *and* $C(x)$ *with loss function* $k(x, u)$. *Let* $k(x, u) \leq k_1(x, u)$ *and*

$$\inf_{x,u} k(x, u) \geq \epsilon > 0.$$

*Then*

$$E_x^w \tau_w \leq V(x)/\epsilon < \infty.$$

Proof.

Let $\tau$ be any random time with $E_x^u \tau < \infty$. The first statement follows from

$$V(x) - E_x^u V(x_\tau) = E_x^u \int_0^\tau k_1(x, u) dt \geq \epsilon\, E_x^u \tau$$

and from the boundedness of $V(x)$ in $X$. (If $E_x^u \tau_u = \infty$, we could increase $\tau$ until $\epsilon\, E_x^u \tau > V(x)$ in $X$.)

Now, by Theorem 1,

$$V(x) \geq \underline{C}(x) = E_x^w \int_0^{\tau_w} k(x, u) \geq \epsilon E_x^w \tau_w$$

and the second statement follows. The existence of a $\tau_w$ is part of the statement on the existence of an optimum $w$.

Theorem 4 gives a method of selecting $S$ so that the corresponding problem can be studied by means of Liapunov functions.

Theorem 4

Let $u$ be admissible and let $V(x)$ be in $D(L)$. Define the sets $R^u = \{x : L^u V(x) \geq 0\}$ and $S_\gamma = \{x : V(x) \leq \gamma\}$.

Let the sets be non-empty and let $R^u$ be a proper subset of $S_\gamma$.

Let $x_0 = x$ be in $X - S_\gamma$ and define $\tau_u$ as the random time of arrival at $\partial S_\gamma$. Then

(25) $$E_x^u \tau_u < \infty.$$

(26) $$P[\sup_{\tau_u \geq t \geq 0} V(x_t) - \gamma \geq \lambda] \leq E[V(x) - \gamma]/\lambda$$

If $w$ minimizes $L^u V(x)$ in $X - S_\gamma$, and $L^w V(x) = -k(x) \leq -\epsilon < 0$ in $X - S_\gamma$, then $w$ is the optimal control for the loss $k(x)$ and target set $S_\gamma$. The cost is

(27) $$\underline{C}(x) = V(x) - \gamma = E_x^w \int_0^{\tau_w} k(x)dt.$$

Also, if $L^u V(x) + k(x) \leq 0$ in $X - S_\gamma$, then

(28) $$V(x) - \gamma \geq C^u(x) = E_x^u \int_0^{\tau_u} k(x)dt.$$

Proof:

Since $L^u V(x)$ is continuous, and $R^u$ is a proper subset of $S_\gamma$, $L^u V(x) \leq -\epsilon < 0$, for some $\epsilon$, in $X - S_\gamma$. Consequently (25) follows from Theorem 3.

Since $V(x) - \gamma \geq 0$ in $X - S_\gamma$ and $L^u[V(x) - \gamma] \leq 0$ in $X - S_\gamma$, (26) follows from (10). The fact that the $u$ which minimizes $L^u V(x)$ is an optimal control for loss $k(x) = -\min_u L^u V(x)$ and target $\partial S_\gamma$ follows from Corollary 1. (27) and (28) follow from Theorem 1 and Corollary 1, by using $V(x) - \gamma$ in their proofs.

Discussion:

For a given Liapunov function $V(x)$, the control problem may be studied in several ways.

Let the loss be $k(x, u)$. Now compute $R^u = \{x : L^u V \geq 0\}$. Now choose a $\gamma$ such that $S_\gamma \supset R_0^u$ and check that $X - S_\gamma$ is not empty. Check that $L^u V(x) + k(x, u) \leq 0$ in $X - S_\gamma$. Then, Theorem 4 says that, starting from a point in $X - S_\gamma$, the total cost, $C^u(x)$ of transferring $x_0 = x$ to a point on $\partial S_\gamma$ is no greater than $V(x) - \gamma$. If $L^u V(x) + k(x, u) = 0$, then the cost is $V(x) - \gamma = C^u(x)$.

Now let $u_1$ and $u$ be given and check that $S_\gamma \supset (R^u \cup R^{u_1})$ and that $X - S_\gamma$ is not empty. If $L^u V(x) + k(x, u) \leq L^{u_1} V(x) + k(x, u_1) = 0$, then the theorem says that the cost of transferring $x = x_0$ in $X - S_\gamma$ to $\partial S_\gamma$ is no greater with $u$ than with $u_1$. Theorem 2 may be used to try to find improved controls, provided that $V(x)$ and $k(x)$ are given.

If, for some $S$, two Liapunov functions $V_1(x)$ and $V_2(x)$ are given with the properties $V_i(\partial S) = \gamma_i$, and $L^u V_1(x) + k(x, u) \leq 0$, $L^u V_2(x) + k(x, u) \geq 0$, $L^u V_i(x) < 0$ in $X - S$, then the cost of transferring $x_o = x$ to $\partial S$ is bounded by

$$V_2(x) - \gamma_2 \leq C^u(x) \leq V_1(x) - \gamma_1$$

Obviously, the cost of transferring to a point interior to $S$ is no less than the cost of transferring to $S$ (by the continuity of $x_t$). The cost of transferring to a set enclosing $S$ is no greater than the cost of transferring to $S$. The observation yields bounds for terminal sets other than the $S$.

Other forms of boundary conditions and loss functions and possible (for example, in case of instability we may minimize the probability of being lost) and will be considered in the examples. Choosing suitable $V(x)$ is, of course, no easier in the stochastic case than in the deterministic case. We have the double problem of finding $V(x)$ so that both $k(x, u)$ and $S$ are suitable.

In the (homogeneous) deterministic case when $\dot{V}(x) \leq 0$ with equality implying $\dot{x} = 0$, it is possible to transfer $x_o = x$ to the origin. This is possible in the stochastic case (with probability one) if $L^u V(x) \leq 0$ with equality only when $x = 0$.

The following theorem is useful for obtaining probability bounds on the rate of convergence of $x_t$ to $\partial S$. The quantity $\alpha$ may depend on the control.

Theorem 5

*Let* $V^n(x)$ *be in* $D(L)$ *and*

$$L^u V^n(x) \le - \alpha V^n(x), \quad \alpha > 0,$$

*in* $X - S$. *Let* $t(\tau) = \min[t, \tau]$, *for* $t$ *non-random. Then*

$$P[\sup_{\tau_u \ge t \ge s(\tau_u)} V(x_t) \ge \lambda] \le e^{-\alpha s} V^n(x)/\lambda^n,$$

*where* $x = x_0$.

*Proof*

Modify the system in $S$ only, so that $L^u V(x) \le - \alpha V(x)$ in $X$. Let $x'$ be the modified trajectory and $\tau'$ the time to the origin for the modified trajectory. By the continuity of the paths

$$P[\sup_{\tau' \ge t \ge s(\tau')} V^n(x_t') \ge \lambda^n] \ge P[\sup_{\tau_u \ge t \ge s(\tau_u)} V^n(x_t) \ge \lambda^n].$$

By Theorem 5 of [15], the left side is less than $e^{-s\alpha} V^n(x)/\lambda^n$, if $x_0 = x$, and the proof is concluded.

## III. EXAMPLES

*Example* 1. Let†

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= (-x - x_2 + u)dt + \sigma(x)dz \end{aligned} \tag{29}$$

with

† The spaces $X$ of the examples are not compact. However, by letting $\sigma^2 = 0$ for large $\|x\|$, and confining $x_0 = x$ to some large, but compact, set, the space may be compactified with little loss in generality.

$$\sigma^2(x) = x_1^2\,\sigma^2, \quad \sigma^2 < 2$$
$$k(x,\ u) = x_1^2 + x_2^2 + u^2\ .$$

If $\sigma^2 < 2$ and $u = 0$, then $x_t \to 0$ with probability one [15]. Owing to this, the target set may be the origin. Theorem 2 will be applied to the computation of a control. With $u = 0$, there is a positive definite quadratic form [15]

$$V(x) = b_{11}x_1^2 + 2b_{12}x_2x_1 + b_{22}x_2^2$$
$$b_{11} = \tfrac{1}{2} + (2 + \sigma^2)/(2 - \sigma^2)$$
$$b_{12} = \tfrac{1}{2} + \sigma^2/(2 - \sigma^2)$$
$$b_{22} = 2/(2 - \sigma^2)$$

so that

$$L^0V(x) = -\,k(x,\ 0) = -\,x_1^2 - x_2^2$$
$$C^0(x) = E_x^0 \int_0^{\tau_0}(x_1^2 + x_2^2)dt = V(x).$$

By Theorem 2 (Eq. (24)), for any $u$ such that

(30) $$u(\partial V/\partial x_2) + u^2 < 0,$$

we have

(31) $$C^u(x) - C^0(x) < 0.$$

In particular,

$$u = -(\partial V/\partial x_2)/2 = -(b_{11}x_1 + b_{22}x_2) \tag{32}$$

satisfies (30). Although the improvement (31) is difficult to estimate, an estimate of the stability improvement may be obtained with the use of (11). Now, with (32), and any real number $n \geq 1$,

$$\begin{aligned} L^u V^n &= nV^{n-1}\left[\frac{\partial V}{\partial x_1} x_2 + \frac{\partial V}{\partial x_2}(-x_1 - x_2 + u)\right] \\ &\quad + \frac{\sigma^2 x_1^2}{2}\left[nV^{n-1}\frac{\partial^2 V}{\partial x_2^2} + n(n-1)V^{n-2}\left(\frac{\partial V}{\partial x_2}\right)^2\right] \\ &= nV^{n-1}[-x_1^2 - x_2^2 - 2(b_{12}x_1 + b_{22}x_2)^2 + 2(n-1)\sigma^2(b_{12}x_1 + b_{22}x_2)^2(x_1^2/V)]. \end{aligned} \tag{33}$$

The middle entry, $-2(b_{12}x_1 + b_{22}x_2)^2$, is due to the control. The noise contributes to all other terms. Also $(x_1^2/V) \leq (b_{11} - b_{12}^2/b_{22})^{-1}$.

Since the control contribution is proportional to the term containing $\sigma^2$, and is of opposite sign, some cancellation occurs; this cancellation increases the maximum value of $n$ for which $L^u V^n$ is non positive in $X$. As $n$ increases, the estimate

$$P[\sup_{\tau_u \geq t \geq 0} V(x_t) \geq \lambda] \leq V^n(x)/\lambda^n \tag{34}$$

improves. With properly chosen $V(x)$, estimates of the form of (34) can yield useful information on the effect of the particular controls. By Theorem 3, if $S$ is a set containing the origin as an interior point, then the average time to $S$ is finite. In any case, $x_t \to 0$ w.p.1.

A useful general form is

$$L^uV^n = nV^{n-1}[L^oV + (L^u - L^o)V + (n-1)\sum_{i,j} \frac{(\partial V/\partial x_i)(\partial V/\partial x_i)}{2V} S_{ij}]$$

$$S_{ij} = \sum_k \sigma_{ik}\,\sigma_{jk}$$

In our case,

$$(L^u - L^o)V = u(\partial V/\partial x_2) = 2(b_{12}x_1 + b_{21}x_2)u$$

Let numbers $\delta > 0$ and $\epsilon < \lambda$ be given, and assume that $x_o = x$ is in $\{x: V(x) \le \epsilon\}$. We will compute the 'smallest' control which guarantees (according to our estimates and method) that $x_t \to 0$ w.p.1 and

$$P[\sup_{\tau_u \ge t \ge 0} V(x_t) \ge \lambda] \le \delta$$

First compute the least $n \ge 1$ such that

$$\sup_x V^n(x)/\lambda^n = (\epsilon/\lambda)^n = \delta.$$

Then $L^uV^n \le 0$ in X if

$$(35) \qquad -x_1^2 - x_2^2 + 2u(b_{12}x_1 + b_{22}x_2) + 2(n-1)\sigma^2(b_{12}x_1 + b_{22}x_2)^2(x_1^2/V) \le 0.$$

A suitable control can be determined from

$$2u(b_{12}x_1 + b_{22}x_2) = \min[0,\ x_1^2 + x_2^2 - 2(n-1)\sigma^2(b_{12}x_1 + b_{22}x_2)^2(x_1^2/V)],$$

which always yields a bounded control (in any compact set) (if $b_{12}x_1 + b_{22}x_2 = 0$, then $u = 0$).

Example 2. Same as Example 1, but let $\sigma^2(x) = \sigma^2$, a constant. We would prefer a $V(x)$ such that $L^o V(x) = -k(x, 0)$. Not being able to find such a $V(x)$, we select one which yields an approximation. If

$$V(x) = 3/2\, x_1^2 + x_1 x_2 + x_2^2 ,$$

then

$$L^o V(x) = -x_1^2 - x_2^2 + \sigma^2.$$

To satisfy the conditions of Theorem 4, let

$$S \supset \{x: x_1^2 + x_2^2 \leq \sigma^2\} = R^o .$$

$L^o V(x) < 0$ in the complement of $R^o$. Although $R^u$ can be made smaller than $R^o$, the minimum eigenvalue will be the same, and the allowable reduction in the size of the target set may not be appreciable. Following the procedure of Theorem 4,

$$C^o(x) = E_x^o \int_0^{\tau_o} (x_1^2 + x_2^2 - \sigma^2)\, dt = V(x) - \gamma, \tag{36}$$

where $\tau_o$ is the random time to the assumed target set

$$S = \{x: V(x) \leq \lambda\} \supset R^o.$$

$C^u(x) < C^o(x)$ if

$$u^2 + (\partial V/\partial x_2)u < 0,$$

which is satisfied (and is minimum) if

$$u = -(x_1/2 + x_2). \tag{37}$$

Also

$$L^u V^n(x) = nV^{n-1}(x)[-x_1^2 - x_2^2 + \sigma^2 - 2(x_1/2 + x_2)^2 + 2\sigma^2(n-1)(x_1/2 + x_2)^2/V(x)],$$

The $-2(x_1/2 + x_2)^2$ term is contributed by the control (37). As in example 1, the control improves stability--in the sense that the probability of an arbitrary increase in $V(x_t)$ (before absorption on $\partial S$) is decreased.

The method may be used to obtain bounds on moments.

Replace $\tau_o$ in (36) by a non-random variable $t$, let $\sigma^2 = 0$ for very large $\|x\|$ (so that $X$ is compact), and assume that each $E\, x_i^2$ converges to a constant as $t \to \infty$. Since $x_1^2 + x_2^2$ is bounded in $X$, the order of integration may be changed for any finite $t$. Then, (36) and the boundedness of $V(x)$ in $X$ yield that

$$\lim_{t\to\infty} E(x_1^2 + x_2^2) \leq \sigma^2$$

and the limit converges to $\sigma^2$ as the point $\|x\|$ of truncation of $\sigma^2$ goes to infinity.

Example 3. Assume the system of example 2. We consider another type of criteria by which $V(x)$ may be chosen. Let $x_o = x = (0, x_{20})$. Then

$V(x_o) = b_{22}x_{20}^2$. Find a $u$ which will transfer $x_2^2$ to some small value $\beta^2$ and such that, for a given $\delta$ and $\epsilon > x_{20}^2$,

$$P[\sup_{\tau_u \geq t \geq 0} x_{2t}^2 \geq \epsilon] \leq \delta.$$

Let $L^u V^n(x) < 0$ for $x_2^2 > \beta^2$, and let $\tau_u$ be well-defined.

Any quadratic form in two variables may be written as

$$V(x) = b' x_2^2 + (b_{11}x_1 + b_{12}x_2)^2 / b_{11} \tag{38}$$

$$b' = (b_{22} - b_{12}^2/b_{11}),$$

where the first term of $V(x)$ is positive definite, and the second is positive semi-definite. Since

$$P[\sup_{\tau_u \geq t \geq 0} b' x_2^2 \geq \lambda] \leq P[\sup_{\tau_u \geq t \geq 0} V(x) \geq \lambda] \leq (b_{22}x_{20}^2)^n/\lambda^n, \tag{39}$$

where $\lambda/b' = \epsilon$, it seems reasonable to use the positive definite quadratic form with the maximum value of

$$b'/b_{22} = 1 - b_{12}^2/b_{11}b_{22},$$

provided, of course, that

$$L^u V^n(x) < 0 \quad \text{for} \quad x_2^2 > \beta^2 \tag{40}$$

and a suitable $n$. The problem suggests that we seek a $V(x)$ such that $L^o V(x) = -x_2^2 +$ constant. Thus, let

$$V(x) = (x_1^2 + x_2^2)/2,$$
$$L^u V(x) = -x_2^2 + \sigma^2/2 + ux_2.$$

If $\beta^2 < \sigma^2/2$, then the use of

$$u = -x_2(\sigma^2/2 - \beta^2)/\beta^2, \quad \beta^2 \leq x_2^2 \leq \sigma^2/2 \tag{41}$$
$$= 0 \qquad \text{otherwise}$$

assures that (40) is satisfied for $n = 1$. Thus, there is a control for which $x_2^2 = \beta^2$ is attainable. Note also that $b'/b_{22}$ is maximum. If $b'/b_{22}$ were not maximum, then either some systematic procedure for maximization would be followed, or else several $V(x)$ would be tried and compared.

To complete the analysis, find the least $n \geq 1$ for which

$$(b_{22}x_{20}^2/\epsilon b')^n = \delta$$

and choose the most convenient $u$ for which

$$L^u V^n(x) = nV^{n-1}[-x_2^2 + \sigma^2 + \frac{(n-1)\sigma^2 x_2^2}{(x_1^2 + x_2^2)} + ux_2]$$

is negative in the desired region $x_2^2 \geq \beta^2$.

There are, of course, similar procedures for more general initial conditions. The quadratic forms may be chosen by selecting the non-constant, non-positive quadratic part of $L^0V(x)$, and solving for $V(x)$.

Other forms of experimentation with the type of quadratic form is possible; e.g., choose a control first (say, of an arbitrary linear form with coefficient to be determined), then choose $x'Bx$, so that the target

$\{x: x'Bx \in \lambda\}$ is of some useful shape, and, finally, compute the contrcl coefficients.

Remark:

Generally, the Liapunov functions $V^n(x)$ do not give the best probability bounds on, say, the excursions of some component $|x_i|$, since it couples the effects of the various components of $x$ more than is necessary. For example, instead of choosing $n = 2$, a suitably chosen homogeneous positive-definite quartic form will usually yield better estimates on the probabilistic behavior is being investigated. The powers of the quadratic form are used here purely for numerical simplicity.

Example 4. Let

$$dx = (Ax + Cu)dt + \sigma dz$$

$$k(x,u) = F(x) + g(u)$$

$$F(x) = \sum_1^n F_{2i}(x),$$

where $F_{2i}(x)$ is a homogeneous positive definite form of order $2i$, and $A$ is stable. By a theorem of Liapunov [18], if $\sigma = 0$, and $u = 0$, there is a homogeneous positive definite function $V_{2i}(x)$ of $2i$-th order, with $\dot{V}_{2i}(x) = -F_{2i}(x)$. When $\sigma$ is a constant matrix not identically zero,

$$L^o V_{2i}(x) = -F_{2i}(x) + Q_{2(i-1)}(x)$$

$$Q_{2(i-1)}(x) = \frac{1}{2} \sum_{j,m} \frac{\partial^2 V_{2i}}{\partial x_j \partial x_m} S_{jm}$$

$$Q_o = \text{constant}.$$

$$S_{jm} = \sum_i \sigma_{ji}\sigma_{mi}$$

$Q_{2i}$, $i \neq 0$, is a homogeneous non-negative definite form of order $2i$.
A Liapunov function

$$V(x) = \sum_1^n V_{2i} \tag{42}$$

with

$$L^o V(x) = -F(x) + Q_o \tag{43}$$

is easily determined: set $\sigma = 0$ and solve, by Liapunov's theorem,

$$\dot{V}_{2n}(x) = -F_{2n}(x),$$

and, in general, for the case $0 < i < n$,

$$\dot{V}_{2i} = -F_{2i}(x) - Q_{2i}(x).$$

If the target set $S = \{x: V(x) = \gamma\}$ includes $\{x: F(x) \geq Q_o\}$, then

$$C^o(x) = E_x^o \int_o^{\tau_o}(F(x) - Q_o)dt = V(x) - \gamma.$$

By Theorem 2, if

$$L^u V(x) + g(u) < L^o V(x),$$

then

$$C^u(x) = E_x^u \int_0^{\tau_u} (F(x) + g(u) - Q_o)dt < C^o(x).$$

For the deterministic problem, this approach was investigated in considerably more detail in [8].

Example 5. Let

$$\text{(44)} \qquad \begin{aligned} dx &= (Ax + u)dt + \sigma dz \\ k(x, u) &= -\rho \\ \dot{A}' + A &= 0, \quad \sigma_{ij} = \sigma^2\delta_{ij}, \end{aligned}$$

and $u'u = \rho^2$. The target set is to be a sphere about the origin, with radius $r > 0$. The deterministic part of (44) has been termed 'norm-invariant'. Let the components of $z$ be independent. The Liapunov function which is the minimal cost of transferring $x = x_o$ to the origin, for the deterministic problem, is

$$V_1(x) = \|x\| = (x'x)^{1/2} .$$

We have

$$L^u V_1(x) = \frac{x'u}{\|x\|} + \frac{\sigma^2}{2\|x\|} ,$$

which is minimized by

$$\text{(45)} \qquad u = -\rho x/\|x\| ,$$

the optimal deterministic control, and

$$L^u V_1(x) = -\rho + (s-1)\sigma^2/2\|x\|, \tag{46}$$

where $s$ is the dimension of $x$. If the target set has a radius at least $\sigma^2/2\rho$, then Theorem 1 yields

$$V_1(x) - \gamma < \underline{C}(x) \ ,$$

The fact that (46) is still a function of $\|x\|$ suggests that $\underline{C}(x)$ is a function of $\|x\|$. Let us try

$$V(x) = \|x\| + a \log x'x + c, \tag{47}$$

where $a$ and $c$ are constants. (47) is suggested by the form of (46). (It is also suggested by the observation that the 'deterministic' contribution to $L^u V$, of $\log x'x$, $2Ax'/x'x$, is of the proper form to cancel part of the 'stochastic' contribution of $V_1(x)$ to $L^u V$, which is $(s-1)\sigma^2/2\|x\|$.)

$$L^u V(x) = \frac{x'u}{\|x\|} + \frac{\sigma^2(s-1)}{2\|x\|} + \frac{ax'u}{x'x} + \frac{a\sigma^2(s-2)}{x'x} \ .$$

With (45),

$$L^u V(x) = -\rho - \frac{a\rho}{\|x\|} + \frac{\sigma^2(s-1)}{2\|x\|} + \frac{a\sigma^2}{x'x}(s-2).$$

Let $s = 2$ and $a = \sigma^2/2\rho$, then $L^u V(x) = -\rho$. At the target set boundary,

$\partial S = \{x\colon x'x = r^2\}$, we have $V(x) = 0$. Thus, for arbitrary $r > 0$, set

$$c = -r - a\log(r^2).$$

Now, $V(x) > 0$ and $L^u V(x) = -\rho$ in $X - S$, and

$$C^u(x) = V(x) = E_x^u \int_0^{\tau_u} \rho\, d\tau = E_x^u \tau_u.$$

Also, since (45) minimizes $L^u V(x)$ over all admissible controls, by Corollary 1, (45) is the average-time-optimal control over the class of admissible controls. If $s > 2$, the procedure may be repeated. This will be developed elsewhere.

Example 6. Take the scalar case

$$dx = -xdt + udt + \sigma dz$$

$$|u| \le 1, \quad S = \{0\},$$

$$C^u(x) = E_x^u \int_0^{\tau_u} (k + |u|)dt.$$

The optimal deterministic solution is (the deterministic version is a problem in [21])

$$\text{(48)} \qquad C_d(x) = \begin{cases} V'(x) = (k+1)\log(|x| + 1), & |x| < k, \quad u = -\operatorname{sign} x \\ V''(x) = (k+1)\dfrac{\log(k+1)\log|x|}{\log k}, & |x| \ge k, \quad u = 0. \end{cases}$$

$$\text{(49)} \qquad \begin{aligned} L^u V'(x) &= -(k+1) - \sigma^2(k+1)/2(|x| + 1)^2, \quad u = -\operatorname{sign} x \\ L^u V''(x) &= -k - \sigma^2 k/x^2, \quad u = 0. \end{aligned}$$

# ERRATA

Stochastic Stability and the Design of Feedback Controls

H. J. Kushner

C.D.S. T.R. 65-5

| Page | | | |
|---|---|---|---|
| 26 | 6th line of Remark | : | the probabilistic behavior. The powers of the quadratic form |
| 29 | 3rd line below (47) | : | Replace $2Ax'/x'x$ by $-2\rho/\lVert x\rVert$. |

At $|x| = k$, $\underline{C}_d(x)$ does not have a derivative. This is not important in the scalar case. (It can be assumed that $\sigma^2(x)$ satisfies (2) and is zero in a small neighborhood of $|x| = k$, with an insignificant change in the process.) Since $L^u\underline{C}_d(x) < -k(x, u)$ for $|x| \neq k$, Theorem 1 yields

$$C^u(x) < \underline{C}_d(x) .$$

The loss for the stochastic problem is less than $\underline{C}_d(x)$, since the problem is scalar and $\underline{C}_d(x)$ is convex downward. Such an improvement is uncommon for vector problems.

By Corollary 1, if $V(x)$ and u satisfied $L^uV(x) + k + |u| = 0$ and $L^{u'}V(x) + k + |u'| \geq 0$ for $u' \neq u$, then u is an optimal control and $V(x) = \underline{C}(x)$. Then, u must satisfy

$$\begin{aligned} (50) \qquad & u = -\operatorname{sign} dV/dx, \quad dV/dx > 1 \\ & u = 0 , \qquad\qquad\quad \text{otherwise,} \end{aligned}$$

exactly the <u>form</u> of the deterministic optimal control. The form (50) is not admissible, but may be approximated arbitrarily closely by an admissible control. Since the problem can be well defined and solvable with a slight modification of $\sigma^2(x)$, the inadmissibility will be ignored. Since $C^u(x) < V'(x)$, $|x| < k$, it is suggested that $|d\underline{C}(x)/dx| < |dV'(x)/dx|$, $|x| < k$, and, hence, that the optimal control would be of the form

$$\begin{aligned} & u = -\operatorname{sign} x, \quad |x| < k' < k \\ & u = 0 , \qquad\qquad |x| \geq k' . \end{aligned}$$

The qualitative information inferred above can be substantiated by solving the exact stochastic problem (which is easy and will not be done here).

Define $S_\gamma = \{x: |x| \leq \gamma\}$, $\gamma > 0$ and let $X$ be a large set containing the origin with $\sigma^2 = 0$ outside $X$. Now, since $L^0 \log(1 + |x|) < 0$ in $X - S_\gamma$ for any $\gamma > 0$, (10) yields

$$P[\sup_{\tau_0 \geq t \geq 0} \log(1 + |x_t|) \geq \lambda] \leq \log(1 + |x|)/\lambda.$$

Better bounds can be obtained if $S_\gamma$ is more restricted. Let $V(x) = |x|^n$, $n \geq 2$. Then

$$L^u V(x) = n|x|^{n-2} A(x)$$
$$A(x) = -x^2 + ux + (n-1)\sigma^2/2.$$

If $A(x) < 0$ in $X - S_\gamma$, then

$$P[\sup_{\tau_u \geq t \geq 0} |x_t| \geq \lambda] \leq |x|^n/\lambda^n .$$

The smallest $S_\gamma$ (such that $L^u V < 0$ in $X - S_\gamma$ and $|u| \leq 1$) corresponds to

$$\gamma = \gamma_1 = [-1 + (1 + 2(n-1)\sigma^2)^{1/2}]/2 \quad ,$$

and then we require $u = -\operatorname{sign} x$ for $\gamma_1 \leq |x| \leq \gamma_0$, where $\gamma_0^2 = (n-1)\sigma^2/2$.

REFERENCES

1. R. W. Bass, discussion of paper by A. Letov in Proc. Heidelberg Conf. on Automatic Control, R. Oldenbourg, Munich (1957), pp. 209-210.

2. R. E. Kalman, S. E. Bertram, "Control System Analysis and Design via the Second Method ov Liapunov," J. Basic Eng., 82 (ASME), pp. 371-393 (1960).

3. J. P. LaSalle, "Stability and Control," J. SIAM Control, Ser. A, 1, no. 1, pp. 3-15 (1962).

4. G. Geiss, "The Analysis and Design of Nonlinear Control Systems via Liapunov's Second Method," Rpt. No. RTD-TDR-63-4076, Grumman Aircraft Corp., Bethpage L. I., (1964).

5. N. E. Nahi, "On the Design of Optimal Systems via the Second Method of Liapunov," AC-9, pp. 274-275, (1964).

6. Z. V. Rekasius, "Suboptimal Design of Intentionally Nonlinear Controllers," IEEE Trans. on Automatic Control, AC-9, pp. 380-385 (1964).

7. G. W. Johnson, "Synthesis of Control Systems with Stability Constraints via the Second Method of Liapunov." IEEE Trans. on Automatic Control, AC-9, pp. 270-273 (1964).

8. R. W. Bass, R. F. Weber, "On the Suboptimal Control of Autonomous Linear Systems," Hughes Aircraft Corp., Rept., March, 1965.

9. J. L. Doob, Stochastic Processes, Wiley, New York, 1953.

10. E. B. Dynkin, Markov Processes, Springer-Verlog, (1965).

11. E. Wong, M. Zakai, "On the Relationship Between Ordinary and Stochastic Differential Equations," Report 64-26, Electronics Research Lab., Univ. of California, Berkeley, Cal. (1964).

12. K. Ito, "Lectures on Stochastic Processes," Tata Institute, Bombay, 1961.

13. E. B. Dynkin, Foundations of the Theory of Markov Processes, Springer-Verlog, (1961).

14. H. J. Kushner, "On the Stability of Stochastic Dynamical Systems," Proc. Nat. Acad. of Sci., 53, pp. 8-12, (January, 1965).

15. H. J. Kushner, "On the Theory of Stochastic Stability," Report 65-1, Center for Dynamical Systems, Brown University, (1965).

16. R. Z. Khas'minskii, "On the Stability of the Trajectory of Markov Processes," Appl. Math. and Mech. (PPM), 26, pp. 1554-1565, (1962).

17. W. M. Wonham, "On Weak Stochastic Stability of Systems Perturbed by Noise," Report 65-4, Center for Dynamical Systems, Brown University (1965).

18. I. G. Malkin, Theory of Stability of Motion, AEC Translation No. 3352, Dept. of Commerce, Washington, D. C., (1958).

19. M. Athans, P. Falb, R. T. Lacoss, "Time, Fuel, and Energy-Optimal Control of Nonlinear Norm-Invariant Systems," IEEE Trans. on Automatic Control, AC-8, pp. 196-201 (1963).

20. R. E. Kalman, "Optimal Control and the Calculus of Variations," Research Institute for Advanced Study (RIAS) Report 61-3 (1961).

21. M. Athans, P. Falb, "Optimal Control: An Introduction to the Theory and Its Applications," McGraw-Hill, (to appear in 1965).